

# A clock containing a massive object in a superposition of states; what makes Penrosian wavefunction collapse tick?

Tjerk H. Oosterkamp and Jan Zaanen  
*Leiden Institute of Physics, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands*  
 (Dated: July 29, 2014)

PACS numbers: 04.20.Cv, 07.10.Cm, 85.70.-w, 03.65.Ta

Penrose has been advocating the view that the collapse of the wave function is rooted in the incompatibility between general relativity and quantum mechanics. Asserting that gravity wins, he arrived on basis of dimensional analysis at an estimate for the collapse time. This is an ad-hoc affair and here we will present an analysis revolving around the explicit role of time in gravitational wave function collapse. We present a thought experiment, in which we investigate the behavior of a hypothetical clock containing a component which can be in a superposition of states. The clocking action comes from a single strong laser pulse traveling in an optical cavity. At its center it contains a massive object, whose only purpose is to introduce a curvature of space time into the problem. We find that a state of this massive object with a smaller radius, but with the same mass, experiences a larger time delay. Considering a coherent superposition of the large and small object, this suffers from an ambiguity in the definition of a common time for both states. We assert that this time ambiguity will affect the relative phase and that the wave function collapse will occur when this extra phase becomes of order unity. An absolute energy scale enters this equation and we recover Penrose's estimate for the collapse time when we set this equal to the rest mass of the object. This sheds light on the counterintuitive aspect of Penrose's claim that gravitational time dilation effects can play any role dealing with masses of order of micrograms: the rest mass is sufficiently large such that the tiny time differences can confuse the unitary time evolution.

## I. INTRODUCTION.

The plan of this paper is as follows. Penrose [1] has put forward an approach to estimate the time scale at which gravity will start to play a role in the quantum mechanical time evolution of heavy objects, which we discuss in detail in section II. Conceptually this revolves around gravitational time dilations which are unequal for the Schrödinger cat states, supposedly eventually destroying

their coherent superposition. However, in Penrose's dimensional analysis the precise role of these time dilations is implicit. Utilizing Einstein's heuristic principle of a clock, in section III we present a gedanken experiment tailored to make it easy to track the progress of time explicitly, inspired on a famous GR experiment based on an analysis by Shapiro. In section IV we present a very simple consideration demonstrating how this interferes with the unitary time evolution of a coherent superposition, arriving at a criterium showing when this Einsteinian ambiguity of time becomes noticeable. This turns out to require a quantity with an absolute dimension of energy and taking the rest mass of the object for this, we recover Penrose's dimensional estimate for the collapse time. In section V we refine the argument by considering continuous mass distributions, also highlighting the role of Penrose's assertion that one state in the coherent superposition encounters a clock set by the space time curvature of its quantum partner.

## II. THE PENROSE VIEW ON WAVEFUNCTION COLLAPSE.

The precise status of the collapse of the wavefunction in a quantum measurement has been a highly contentious and confusing subject since its introduction some 80 years ago. A majority view is to assert that there is in fact no problem at all, invoking the "for all practical purposes" idea that conventional decoherence due to the interaction with the environment suffices. However, there is still quite some dissent in the form of a variety of ideas claiming that the collapse cannot be explained on this basis. This invokes alternative interpretations varying from quite mystical (e.g., that human consciousness is the culprit) up to the quite practical "objective collapse" ideas. The latter assert that there is just new physics at work, of a kind that is in principle measurable, while it can eventually be comprehended as a reasonable physical process which does not invoke human observers, many worlds, whatever.

This school of thought has, at the least, the benefit of a prediction – it is not mere philosophy. It is an empirical fact that microscopic objects like electrons or quarks fully submit to the unitary world of orthodox quantum physics, while macroscopic things like cars are

never found in coherent superposition. Given the hypothesis that the collapse is a measurable physical process, it should take place in a regime of scale between the microscopic and the macroscopic. Especially the former has been pushed upwards, by the demonstration that quasi-macroscopic objects like flux qubits [2] do not collapse within the bath-decoherence time scales that can, at present, be achieved in the laboratory. Such an "objective collapse", however, does need to happen when things get quite big.

What determines this scale? While others have suggested to use gravity inspired models [3–5], we believe there is only one proposal available on the basis of known physics: the idea of Penrose that the wave function is based on the incompatibility of the unitary time evolution which is at the heart of quantum physics, and the space time of general relativity [1]. This conflict is manifest in all attempts to get a grip on quantum gravity. The case in point is the Hawking radiation and entropy that follows from combining the vacuum of quantum field theory and the space-time of a Schwarzschild black hole [6]. Upon fixing the frame with a Schwarzschild metric (convenient for an outside observer) one finds out that the coherence of the field theoretical vacuum is "ripped apart" by the black hole horizon, with the effect that the black hole turns into a black body radiator, with a temperature that diverges upon approaching the event horizon. However, in the Lemaître coordinate system of the metric of the observer falling freely into the black hole, the horizon is immaterial and nothing is supposed to happen. It seemed that this apparent paradox could be resolved using the notion of "black hole complementarity" [7], revolving around the idea that it is impossible for the two observers to exchange the information regarding the physics at the horizon, but very recently a flaw was found in this argument in the form of the "firewall paradox" [8].

The modern main stream view is to claim that these troubles are associated with an incomplete description of space-time. Space-time is supposed to be emergent: a classical, coarse grained description of a more general quantum-gravity theory reigning at the Planck scale. Black hole physics is then seen as a way to get information on quantum gravity, asserting that the apparent paradoxes are just revealing that the coarse graining is a subtle affair, repairing GR in this regard. However, there are dissidents (including the Penrose school) claiming that GR is fine but quantum physics is the culprit. Departing from this perspective, the origin of the quantum physics-general relativity incompatibility is quite obvious. GR is controlled by the symmetry principle of general covariance. In the black hole context, this means that the Le Maitre- and Schwarzschild metrics describe the same space time, just differing in the regard that, for reasons of convenience, matters are computed in two different gauge fixes. But the quantum field theory yields

two very different answers: the bottom line is that unitarity, the time evolution governed by linear transformations in Hilbert state with the Hamiltonian as generator of time translations, is not a diffeomorphic invariant.

As stressed by Penrose, unitarity requires a global time like Killing vector [1], and this becomes an issue when gravitationally inequivalent space times are involved in a coherent quantum superposition. In principle, a point like identification between space times with different mass distributions is an impossibility according to general relativity. This in turn is an issue when one is dealing with simple Schrödinger cat states, since the live and the dead cat will have a different mass distribution making it impossible to assign a global time like direction in a space time that can be "shared" by the two cats in the superposition. Henceforth, Schrödinger cats face in principle the same problem as black holes. In the regime of particle physics down to molecular physics, well removed from both the Planck scale and the macroscopic scale, gravity is so weak that this cannot possibly play any role. Therefore, unitarity is just fine in the realms of atoms down to the tera electron volts of the best particle accelerators. But on the human scale, gravity becomes noticeable: could it be that the collapse occurs because gravity wins, destroying unitary evolution and thereby causing the collapse? This is the key question posed by Penrose [1].

At present, any insight in the microscopic theory that would lead to the "gravity wins" outcome is lacking, and in the absence of theoretical guidance, all that remains is dimensional analysis. Penrose suggested that a "Planck scale" can be identified associated with the gravitational wave functional collapse that lies in the regime in between microscopics and macroscopics, based on natural dimensions of quantum physics and gravity. Planck's constant  $\hbar$  is obviously the quantity associated with quantum physics, carrying the dimension of energy times time. This is a convenient dimension to convert energy into time, and Penrose [1] asserts that the time associated with the wave function collapse is given by,

$$\tau_G = \frac{\hbar}{\Sigma_G} \quad (1)$$

where  $\Sigma_G$  is a gravitational quantity with the dimension of energy associated with the inequivalence of space times encountered in Schrödinger cat like situations. He then suggested that this should be the Newtonian gravitational self energy. The cat is surely non-relativistic and in a regime where gravity is weak, and therefore one should look in the Newtonian limit. The "alive cat" defines a gravitational potential well, associated with its mass distribution. The gravitational self-energy is defined by keeping this potential fixed, while one computes the gravitational energy associated with moving the mass distribution to become coincident with its "dead cat"

quantum copy,

$$\begin{aligned}\Sigma_G &= \frac{1}{G} \int d^3\mathbf{x} (\mathbf{f}_a - \mathbf{f}_b) \cdot (\mathbf{f}_a - \mathbf{f}_b) \\ &= 4\pi \int d^3\mathbf{x} [\Phi_a - \Phi_b][\rho_a - \rho_b]\end{aligned}\quad (2)$$

where  $\mathbf{f}_a$  and  $\mathbf{f}_b$  are the vectors indicating the strength of the gravitational fields associated with two different mass distributions,  $\rho_a$  and  $\rho_b$ , that are in superposition with each other.  $\Phi_a$  and  $\Phi_b$  are the gravitational potentials associated with these mass distributions,  $d^3\mathbf{x}$  indicates an integral over the three spatial dimensions and  $G$  is Newton's constant.

Assuming that the cats correspond with simple spherical masses,  $M$ , with radius  $a$  displaced over a length  $d$  where  $a \ll d$ , one arrives at an estimate for the order of magnitude of the gravitational collapse time [1]

$$\tau_G = \frac{5}{6} \frac{\hbar a}{GM^2} \quad (3)$$

Intriguingly, one finds that for a "cat" of typical size  $a = 1 \mu\text{m}$ , which has a weight in the range of micrometer sized bacteria ( $10^{-15} \text{ kg}$ ) and which is in superposition with itself after being displaced by a length of  $b = 1 \mu\text{m}$ ,  $\frac{M^2 G}{a} = 6.6 * 10^{-35} \text{ J}$  and it takes a time  $\tau_G$  of a few seconds to collapse its wavefunction. This is precisely in the range, which has not been explored experimentally. It is however quite appealing for experimentalists, since there is a serious potential that this regime comes into reach using the latest technology, e.g. [9–11]. At the same time, this estimate has been criticised merely on basis that any effect of gravity on time and so forth should be so minute that it can be completely ignored [12]. After all, the dogma is that gravity and quantum physics should only clash at the conventional Planck scale. Our main result is that we will arrive at a rational explanation why this intuition might be in principle misleading.

### III. THE CLOCKS OF SCHRÖDINGER'S CAT: A SHAPIRO TYPE GEDANKEN EXPERIMENT

Even for the purpose of dimensional analysis, Penrose's estimate for the gravitational wave function collapse time is ad-hoc. The assertion that gravity enters via the gravitational self energy, is not rooted in a detailed consideration of how the "ambiguity of time in the superposition" arises in general relativity. Instead, Penrose argues that the gravitational energy is the only quantity he can identify in this Newtonian regime which relates to the superposition of mass distributions, while it can be balanced with  $\hbar$  to yield a reasonable scale.

We wish to point out here that it is in fact quite straightforward to address this ambiguity of time as it

arises in gravity. We employ Einstein's favorite heuristic method of tracking how clocks tick in the reference frames of observers traveling with the Schrödinger's cat quantum copies. Since their mass distributions are different, the clocks attached to the quantum copies will indicate a different time in a classic GR manner, and it is obvious that this disagreement should correspond with the time ambiguity as of relevance to the destruction of the unitary time evolution.

At stake is that, following Penrose, the two quantum copies are characterized by a different sense of time relative to each other. To discern to what extent the sense of time is different in these two space-times, we need a measure that can be shared by both 'universes'. We propose a rather natural clock that can accomplish the goal of measuring the relative difference in the sense of time in this situation. In this paper we estimate the difference in the phase evolution of a mass distribution using two different clocks: one for each reference frame of the two different Schrödinger cat quantum copies. Please note that unlike previous papers, which investigate time dilation effects on quantum mechanics due to a single gravitational potential [13–17] we look at the difference between two time dilation effects of two superposed states. We wish to make it clear however, that we do not derive an expression which provides the new time evolution which ensues in the presence of a superposition of time dilation effects.

In our approach, the gravitational side gives rise to a quantity with the dimension of time (instead of energy as in Penrose's estimate) and we then need to work out how this enters the quantum mechanical equation. We will argue that this involves necessarily an *absolute* energy scale. In section 4, we take for this energy the relativistic rest mass of the "cat" and we find that the result of the dimensional analysis becomes coincident with Penrose's estimate, assuming that we not only take time to depend on the mass distribution of a quantum copy of the cat but that every atom making up the cat can be assigned its own 'clock'.

To start our considerations, we will consider a Gedanken experiment. It is given in by convenience, since it allows us to use material from GR textbooks to estimate the ticking of the clocks. Although we are not aware of any physical principle prohibiting the construction of the device, the barriers to overcome in order to make it work in the laboratory might well be insurmountable. However, we just employ it in order to make the dimensional analysis easy and we expect it to be trustworthy in the limited sense of getting the collapse scale right by order of magnitude.

It consists of a ball made of a material having the property that it undergoes a zero temperature (quantum) first order phase transition where the volume of the material drastically changes. This is less exotic than it might appear. The lanthanides metals cerium, praseodymium and gadolinium as well as the actinide plutonium show a ther-

mal "volume collapse" transition [18] where the volume of these metals can decrease by as much as 15%, as related to a drastic, cooperative change of their f-electron systems from a delocalized- to a localized nature. For recent work in Ce see Ref. [19]. A first challenge for the material scientist is to drive such a transition to very low temperatures. In principle it is possible to force the ball in a coherent superposition of its large- and small volume phases right at the zero temperature transition. Subsequently, the ball has to be kept isolated from the environment in order to prohibit decoherence, while it surely has to be kept at a very low temperature. To accomplish this experimentally could well be an impossible pursuit, but in principle it might be done.

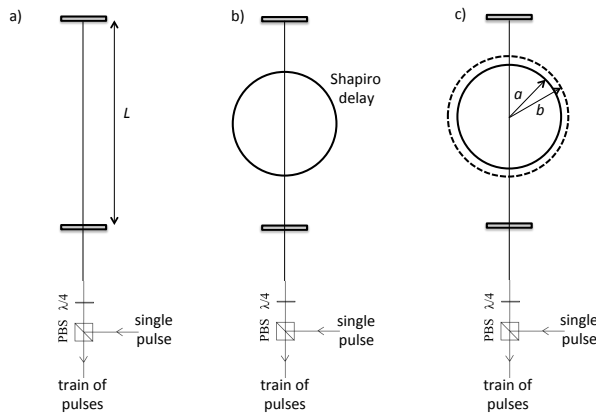


FIG. 1: three clocks, each consisting of a strong laser pulse coupled into a cavity through a polarizing beam splitter and a  $\lambda/4$ -plate, and subsequently bouncing around in a cavity. The lower mirror has a slight transmission, such that a train of pulses is coming out of the cavity. Panel a) and b) show a clock without and with a mass inside the cavity, while panel c) shows a clock in which the mass is in a superposition of states.

Why do we introduce this ball in a superposition of its two "volume states"? The reason is that we can directly apply a famous GR story, as related to the direct measurement of gravitational time dilatation: the Shapiro time delay effects, as tested in the 1960's exploiting the solar system [20]. By reflecting off a radar pulse of the surface of the planet Venus, its time of flight was measured. It turned out that as Venus passes behind the sun, the radio pulse experiences a delay of approximately  $200 \mu\text{s}$ , due to the space time curvature induced by the mass of the sun. One already anticipates that such time dilation effects will be quite delicate when dealing with objects of the weight of E. coli.

For the sole purpose of measuring the times associated with the mass distributions of the large and small ball, we present in Fig. 1 three different clocks. Panel a) shows

a conventional clock, while the other two clocks contain a massive object, which for panel c) is in a superposition of volume states. The conventional clock in panel a) consists of a laser pulse bouncing between two mirrors separated by a distance  $L$ . The end mirror at the top of the image is perfectly reflective, while the entrance mirror is almost perfectly reflective but has a small transmission  $\epsilon$ . The light enters the cavity from a laser  $I$  which emits a single gaussian shaped pulse  $p(t) = \frac{1}{\sqrt{2}} \exp(-t^2/t_{pulse}^2)$  centered around  $t = 0$ ,  $p(t)$  of monochromatic light with wavelength  $\lambda$  and pulse length  $t_{pulse} \ll L/c$ , where  $c$  is the velocity of light. The light exits the cavity as a train of equally spaced pulses,  $p_{clock}(t) = \sum_{i>0} p(t - n\Delta t)$ , separated by  $\Delta t = 2L/c$ . Please note that these light pulses do not necessarily need to be detected.

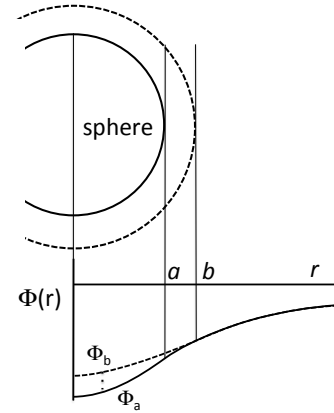


FIG. 2: The gravitational potential near a sphere as a function of the distance from the center of the sphere  $r$ . The solid and dashed curves indicate the potential of a sphere of radius  $a$  and of radius  $b$ , respectively. The dotted vertical line illustrates that, for lack of a theory of quantum gravitation, we don't know the gravitational potential of a sphere in a superposition. The vertical dotted line indicates that the potential at a certain radial position of two states with radius  $a$  and radius  $b$  might be anywhere between the two solid curves.

Panel b) shows a clock, which is influenced by the gravitational field due to the presence of a mass. A spherical mass,  $M$ , resides within the cavity, whose sole purpose is to introduce a space-time curvature, and since the measurement of time is just a Gedanken experiment, the sphere can be regarded as perfectly transparent while it does not affect the laser beam in any way, other than changing the space time curvature. Due to the space-time curvature induced by the sphere, the Shapiro delay [20] is expected to occur, with our spherical mass taking the role of the sun.

In figure 2, we plot the gravitational potential  $\Phi(r)$ , from which the gravitational force,  $F_G$ , felt by a test mass,  $M_{test}$ , can be derived through  $F_G =$

$-M_{test} \frac{d}{dr} \Phi(r)$ . Outside a sphere of radius  $a$  the potential is given by  $\Phi(r) = -GM/r$  with  $G$  Newton's constant and  $r$  the distance from the center of the sphere, while inside the sphere it shows a quadratic behaviour  $\Phi(r) = -(GM/a)(3/2 - r^2/2a^2)$ , since the gravitational force grows linearly inside the sphere. For a light beam running through the center of the sphere, it is straightforward to integrate the gravitational time dilation. The magnitude of this effective time delay, for two passes, from the bottom mirror to the end mirror at the top of the diagram and back,  $\Delta T_a = T_0 - T_a$ , follows from a consideration involving the gravitational index of refraction, which in the radial direction is  $n_G(r) \approx 1 - 2\Phi(r)/c^2$ . Here  $T_0 = \frac{2L}{c}$  is the time for two passes when there is no mass present and  $T_a$ , is the time for two passes when a mass  $M$  with radius  $a$  is present in between the mirrors.

$$\Delta T_a = \frac{2L}{c} - \frac{2}{c} \int_{-L/2}^{L/2} n_G(r) dr \quad (4)$$

$$\Delta T_a = \frac{4}{c^3} \int_{-L/2}^{L/2} \Phi(r) dr \quad (5)$$

This would imply that the ticking of time,  $t_a$ , measured using the period  $T_a$  of our clock with a sphere with radius  $a$  (Fig. 1b), would be slightly slower than the ticking of time,  $t$ , measured using the period  $T_0$ , of the same clock without the sphere (Fig. 1a):

$$T_a = T_0 - \Delta T_a = (1 - \epsilon_a) T_0 \quad (6)$$

$$t_a = (1 - \epsilon_a) t \quad (7)$$

with

$$\epsilon_a = \frac{\Delta T_a}{T_0} = \frac{2}{Lc^2} \int_{-L/2}^{L/2} \Phi_a(r) dr \quad (8)$$

and  $T_0 = 2L/c$ .

Let us now compute the time difference associated with a laser pulse travelling through the two different volume states of the sphere, characterized by the radii  $a$  and  $b$ .

$$\Delta T_a - \Delta T_b = \frac{8}{c^3} \int_0^{L/2} (\Phi_a(r) - \Phi_b(r)) dr \quad (9)$$

$$= \frac{8GM}{c^3} \left( \int_0^a \left[ \frac{3}{2a} - \frac{r^2}{2a^3} \right] dr - \int_0^b \left[ \frac{3}{2b} - \frac{r^2}{2b^3} \right] dr + \int_a^b \frac{dr}{r} \right) \quad (10)$$

and since the first two integrals cancel only the last one remains:

$$\Delta T_a - \Delta T_b = \frac{8GM}{c^3} \ln\left(\frac{b}{a}\right) = \frac{8GM}{c^3} \frac{b-a}{a} \quad (11)$$

where the last step is possible if  $a$  is only slightly smaller than  $b$ .

A sphere with a radius of  $b = 5 \mu\text{m}$ , engineered to have a low temperature volume phase transition of 15% with a density of  $5000 \text{ kg/m}^3$ , will have a typical mass  $M = 3 * 10^{-12} \text{ kg}$ . If a superposition with its low volume state were to be achieved,  $a/b = 0.95$ , and we were to take a distance between the mirrors surrounding the spheres of  $L = 10 \mu\text{m}$ , the time difference of a laser pulse passing these two mass distributions would come out as  $\Delta T_a - \Delta T_b = 2.6 * 10^{-48} \text{ s}$ . and the dimensionless parameter

$$\epsilon_a - \epsilon_b = \frac{\Delta T_a - \Delta T_b}{T_0} = 4 * 10^{-35} \quad (12)$$

This is indeed an exceedingly small effect. It just confirms the intuition that one would not expect relativistic time dilation effects to play any role on the scale of biological cells, given that the sun was barely heavy enough to measure the Shapiro delay using 1960's radar technology.

#### IV. ON THE ABSOLUTE ENERGY OF SCHRÖDINGER'S CAT.

As we have managed to compute quantitatively the difference in time experienced by the two spheres in their different volume states, we now have to find out how this ambiguity in time can affect the coherent superposition. Given that this time ambiguity is a small number, and we have only the modest task of identifying dimensions, we can proceed in a perturbative fashion. Let us first ignore these time delay effects to specify the Schrödinger cat state involving the large- and small volume states of the sphere  $|a\rangle$  and  $|b\rangle$ . We prepare the state at time zero in the superposition  $|\psi(t=0)\rangle = \alpha |a\rangle + \beta |b\rangle$  and allowing for slightly different energies  $E_a$  and  $E_b$  we find that this state will have evolved after a time  $t$  in,

$$|\psi(t)\rangle = \alpha e^{\frac{i}{\hbar} E_a t} |a\rangle + \beta e^{\frac{i}{\hbar} E_b t} |b\rangle \quad (13)$$

We learned in the previous section that because of the gravitational time dilatation effect the two states actually experience a different time, since their gravitational potentials are different! Proceeding naively, pretending that the superposition is still subjected to a unitary evolution, we might instead expect the state,

$$|\psi(t)\rangle = \alpha e^{\frac{i}{\hbar} E_a t_a} |a\rangle + \beta e^{\frac{i}{\hbar} E_b t_b} |b\rangle \quad (14)$$

where  $t_a$  and  $t_b$  are the 'Shapiro' times introduced in Eqn. (7) and computed in the previous section.

This time evolution is not according to the rules that quantum mechanics poses. However, it does express in a minimal way how the quantum time evolution is affected by the ambiguity of relativistic time. Please note that we are merely looking for a time-scale which signals the end of unitary quantum mechanics. Surely the last thing we expect to appear are the familiar Rabi oscillations, since these are rooted in the assertion that the time evolution is unitary. We use the equation above for the purpose of identifying the time-scale after which this might have a manifestation in an experiment; at the moment the phase difference becomes of order unity one is not supposed to use equations of this kind. We therefore proceed by investigating the phase difference

$$\phi(t) = \frac{1}{\hbar}(E_a t_a - E_b t_b) \quad (15)$$

with the two times  $t_a$  and  $t_b$  parametrized in terms of the Shapiro delay parameters  $\epsilon_a$  and  $\epsilon_b$  of the two mass distributions as  $t_a = (1 - \epsilon_a)t$  and  $t_b = (1 - \epsilon_b)t$  associated with the states  $|a\rangle$  and  $|b\rangle$ , respectively. Please note that at the very end of this paper we will investigate a more complicated phase difference

$$\phi(t) = \frac{1}{\hbar}(E_a t_a - E_a t_b + E_b t_b - E_b t_a) \quad (16)$$

and yet other phase differences are thinkable, which we will not consider here.

For simplicity we first investigate Eqn. 15, which can be rewritten as

$$\phi(t) = \frac{1}{\hbar}((E_a - E_b)t - (\epsilon_a E_a - \epsilon_b E_b)t) \quad (17)$$

This consideration reveals an interesting surprise: in the presence of the gravitational ambiguity of time, the time evolution of the wave function *becomes sensitive to the absolute value of the energy of the states involved*. In normal quantum mechanics, i.e. when the two states in superposition agree on their time, only the energy differences  $E_a - E_b$  matter, because the mean energy  $E_{mean} = (E_a + E_b)/2$  appears only in the overall phase  $iE_{mean}t$ , which is therefore pure gauge and devoid of physical implications.

But when the time depends on the state, as in our Shapiro situation, it does matter whether or not  $E_{a,b}$  are replaced by a different energy  $E_{a,b}^*$ , shifted by an offset  $\Delta E$ :

$$E_{a,b}^* = E_{a,b} + \Delta E \quad (18)$$

After this substitution, the phase shift  $\phi(t)$  becomes  $\phi^*(t)$ ,

$$\phi^*(t) = \phi(t) + \frac{t}{\hbar}(\epsilon_a - \epsilon_b)\Delta E \quad (19)$$

We are now facing a question which is unusual in quantum physics, but a quite natural answer is found in general relativity: what to take for the absolute energy  $E^*$ ? Of course the relativist's answer would be  $E^* = Mc^2$ , the rest mass of the sphere. Inserting this and using the expression found for  $\epsilon_a - \epsilon_b$  for our particular geometry, we obtain,

$$\phi^*(t) = \phi(t) + \frac{t}{\hbar} \frac{8GM}{c^2} \frac{b-a}{2La} * Mc^2 \quad (20)$$

$$\phi^*(t) = \frac{t}{\hbar} 8GM^2 \left( \frac{b-a}{2La} \right) \quad (21)$$

where in the last step  $\phi(t)$  has been neglected because  $Mc^2 \gg E_a, E_b$ . Note that the term  $c^2$  in the numerator of the expressions for  $\epsilon_a - \epsilon_b$  has been cancelled by the multiplication with the energy  $Mc^2$ .

If we now insert the values taken for  $L = 10 \mu\text{m}$ ,  $a = 4.75 \mu\text{m}$ ,  $b = 5 \mu\text{m}$  and  $M = 2.6 * 10^{-12} \text{ kg}$  in the previous section, we find that after a time of  $t = 75 \mu\text{s}$  the phase  $\phi^*$  has reached a value of  $2\pi$ . This timescale is very much like the scale that Penrose arrived at with his analysis.

Although our Gedanken experiment departs from a Newtonian, non-relativistic limit we would like to point out that the above expression is invariant under a Lorentz boost. The phase difference discussed here, should not depend on the fact whether the experiment is observed by a spectator which moves in a frame with a velocity  $v$  relative to our 'cat'. This is an issue especially since we have to explicitly invoke the relativistic energy  $E^*$  to arrive at the above expression. The expression  $(\epsilon_a - \epsilon_b)\Delta E = \frac{\Delta T_a - \Delta T_b}{T_0} \Delta E$  is Lorentz invariant because to an outside observer, who sees the clock fly by at a velocity  $v$ , the energy  $\Delta E = Mc^2$  increases to  $\Delta E' = \Delta E / \sqrt{1 - \frac{v^2}{c^2}}$ , while the clock without the masses will appear to tick more slowly,  $T_0$  becoming  $T'_0 = T_0 / \sqrt{1 - \frac{v^2}{c^2}}$ , thereby decreasing the value of  $\epsilon_a - \epsilon_b$ . These two effects exactly cancel because  $T_0$  is in the denominator. Finally, because  $\Delta T_a - \Delta T_b$  does not depend on  $v$ , the expression for  $\phi^*$  is Lorentz invariant.

## V. CONTINUOUS MASS DISTRIBUTIONS: ATOMS ACQUIRING INDIVIDUAL CLOCKS.

Let us now generalize our heuristic clock model: we will argue that our line of thought may lead to the same collapse time as the one proposed by Penrose [1]. While Penrose balances the gravitational self energy with  $\hbar$  to

arrive at an estimate of a characteristic time after which quantum superposition will collapse  $\tau_G = \frac{\hbar}{\Sigma G}$ , we point out that  $\phi^*$  grows with time and that the collapse of the wavefunction might occur by the time  $\phi^* = 2\pi$ , which for our example of the superposition of two states consisting of concentric spheres of different densities, such as described in the previous section, would result in a collapse time

$$\tau_G = \frac{4\pi\hbar La}{8GM^2(b-a)} \quad (22)$$

which indeed contains all the dimensions in the expression one would arrive at when evaluating the integral for the gravitational self energy, called on by Penrose. Of course, the length of our clock does not appear in the integral for the gravitational self energy, nor do we think that the length of our clock should be an ingredient that is fundamental in our analysis.

The way in which we can do away with the length of the clock, allowing us to write our approach in the same integral form as Penrose, is by giving each atom making up the sphere its own clock. This is natural realizing that it is rather unsatisfying that our first model depends on the details of how the clock is constructed, such as the separation of the mirrors  $L$  or whether one chooses to send the light beam right through the center of the sphere or off center. We therefore proceed by giving each atom in the sphere its own clock, to ask subsequently which phase shift the different sphere states pick up when we parametrize the time  $t$  experienced by each separate atom.

We consider a solid body where the system of atoms has spontaneously broken translation symmetry. This implies that the state of the sphere with radius  $a$  ( $|a\rangle$ ) is a product state of all the nuclei in the sphere:

$$|a\rangle = \prod_j |r_j\rangle^a \quad (23)$$

where  $|r_j\rangle$  are real space wave-packets for the  $j$ th nucleus which is part of the sphere localized in its crystal position, and the superscript  $a$  serves as a reminder that all positions  $r_j$  make up the lattice of the sphere with radius  $a$ .

We would like to write down the gradual build up of a phase difference between two states with different mass distributions. We start out with the time evolution of a single state.

$$|a(t)\rangle = \prod_j e^{\frac{i}{\hbar} m_j c^2 t_{j,a}} |r_j\rangle^a = e^{\frac{i}{\hbar} \sum_j m_j c^2 t_{j,a}} \prod_j |r_j\rangle^a \quad (24)$$

where  $m_j$  denotes the mass of the  $j$ th atom. This is perfectly sound when all  $t_{j,a} = t$ . We now deviate from the quantum mechanical rules, however, by asserting that

each atom carries its own time, denoted by the subscripts in  $t_{j,a}$  with

$$t_{j,a} = (1 - \epsilon_{j,a})t \quad (25)$$

with

$$\epsilon_{j,a} = \frac{2\Phi_a(r_{j,a})}{c^2} \quad (26)$$

determined by the gravitational field,  $\Phi_a(r_{j,a})$ , at the position of the  $j$ th atom, which is determined by the configuration of all atoms making up the sphere with radius  $a$ .

Using the same parametrization of the Shapiro like time delay, we arrive at a phase difference,  $\phi_{diff,a-b}$ , between the pure states  $|\psi\rangle = |a\rangle$  and  $|\psi\rangle = |b\rangle$

$$\phi_{diff,a-b} = \frac{1}{\hbar} \left( \sum_j m_j c^2 t_{j,a} - \sum_j m_j c^2 t_{j,b} \right) \quad (27)$$

which becomes, after substituting  $t_{j,a} = (1 - \frac{2\Phi_a(r_{j,a})}{c^2})t$  and  $t_{j,b} = (1 - \frac{2\Phi_b(r_{j,b})}{c^2})t$ , where the extra subscript added to the spatial coordinate  $r_j$  serves as a reminder that each atom has a different coordinate in state  $|a\rangle$  versus state  $|b\rangle$ :

$$\phi_{diff,a-b} = \frac{t}{\hbar} \left[ - \sum_j m_j 2\Phi_a(r_{j,a}) + \sum_j m_j 2\Phi_b(r_{j,b}) \right] \quad (28)$$

which may also be written in integral form as

$$\phi_{diff,a-b} = \frac{t}{\hbar} 2 \int d^3\mathbf{x} [\Phi_b(\mathbf{x})\rho_b(\mathbf{x}) - \Phi_a(\mathbf{x})\rho_a(\mathbf{x})] \quad (29)$$

where  $\rho_a(\mathbf{x})$  and  $\rho_b(\mathbf{x})$  are the mass distributions of the sphere with radius  $a$  and radius  $b$ , respectively.

Please note that it should be understood that the gravitational potential is taken to be smooth at the atomic scale, because we have assumed in our derivation that the gravitational potential is taken to have a single value per atom. In fact this was used to arrive from Eqn. (8) at Eqn.(26).

Now, when asking at which time  $t_G$  does the phase difference accrued between states  $a$  and  $b$  reach  $\phi_{diff,a-b} = 2\pi$ , one arrives at

$$\tau_G = \frac{\hbar}{4\pi \int d^3\mathbf{x} [\Phi_b(\mathbf{x})\rho_b(\mathbf{x}) - \Phi_a(\mathbf{x})\rho_a(\mathbf{x})]} \quad (30)$$

which closely resembles the integral form of the collapse time proposed by Penrose Eq.'s (1,2), containing

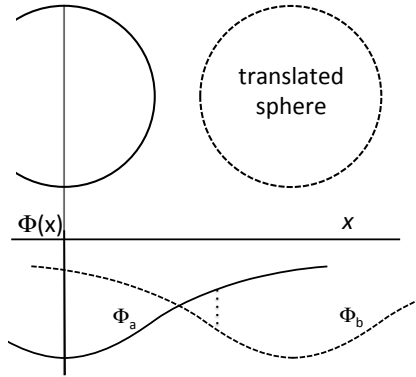


FIG. 3: The gravitational potential near a sphere as a function of position measured along an axis through the center of the sphere,  $x$ . The solid and dashed curves indicate the potential of a sphere of the same radius, but displaced in  $x$ .

two of the four terms in the integral  $\int d^3\mathbf{x}(\Phi - \Phi')(\rho - \rho')$  of Eq. (2).

This is due to the fact that, in the above, we have only calculated the phase difference between the two *separate* product states and we have not tried to answer what the actual time evolution of a superposition of states will be, nor what the gravitational potential would be when it is produced by a superposition of states. For instance, considering the particular case of a sphere in superposition of two states, which are only *displaced* over some distance rather than stretched such as to get a different density (see Fig. 3), the calculated phase difference from equation Eq. (30) would vanish altogether. This is surely the most elementary way to ask the gravitational wave function collapse question in a Schrödinger cat type setting, and we are clearly still missing an ingredient. As pointed out previously, instead of equating the phase difference of Eqn. (15) to  $2\pi$ , one could also choose to equate the phase difference of Eqn. (16) to  $2\pi$ . The reason is actually quite obvious: Penrose has an extra rule wired into his gravitational self-energy construction. In the case of a coherent superposition of states like  $|\psi\rangle = \alpha|a\rangle + \beta|b\rangle$ , it is implicit in the gravitational self energy construction that the time dilation experienced by state  $|a\rangle$  is actually associated with the space time determined by the mass distribution of its quantum copy  $|b\rangle$  and the other way around. Leaving it as a question why this should be the case, it immediately follows that also in the case of the "Shapiro times" the collapse time is associated with the *difference* between the two gravitational fields  $\Phi_a - \Phi_b$ . This yields,

$$t_G = \frac{\hbar}{4\pi \int d^3\mathbf{x}[\Phi_a(\mathbf{x}) - \Phi_b(\mathbf{x})][\rho_a(\mathbf{x}) - \rho_b(\mathbf{x})]} \quad (31)$$

This expression is now exactly the same as Penrose's expression for the characteristic time associated with wavefunction collapse (Eqn. 1,2).

## VI. CONCLUSION.

Surely the power of Penrose's logic is to lift the quantum measurement debate from the philosophical- to the empirical realms. The question whether objective state reduction is the one that is chosen by nature, and the issue whether gravity or quantum physics is loosing out in the real theory of quantum gravity, is only decidable by experiment. All we have accomplished is to construct a simple rational explanation why the ad-hoc dimensional analysis of Penrose can make sense after all. Since the Shapiro experiment was successful, it is evident that our estimate for the gravitational time ambiguity is physical. Also the way that this enters into the time evolution of the quantum superposition is very elementary. Although natural from the point of view of dimensions, the main outcome of our analysis is that in order for the gravitational collapse to happen on the "E. coli scale" the absolute energy scale that is required in quantum mechanics, when time gets ambiguous, has to be the relativistic rest mass. The consequence of a successful gravitational collapse experiment would therefore be that when unitarity comes to an end, the rest mass of Schrödinger's cat is no longer a quantity that can be gauged away. We hope that this will be a guidance for those theorists searching for the "gravity first" theory of quantum gravity.

We thank Jasper van Wezel, Koenraad Schalm, Dirk Bouwmeester, Roger Penrose and Ana Achúcarro for discussions. THO thanks the Foundation for Fundamental Research on Matter (FOM) and the Netherlands Organisation for Scientific Research (NWO) for financial support.

- 
- [1] Penrose, R. *Gen. Rel. Grav.* **28**, 581, (1996); Penrose, R. *Phil. Trans. R. Soc. Lond. A* **356**, 1927, (1998).
  - [2] I. Chiorescu, Y. Nakamura, C. Harmans et al., *Science* **299**, 1869, (2003).
  - [3] F. Karolyhazy. Gravitation and quantum mechanics of macroscopic objects. *Nuovo Chimento A*, 42:390, (1966).
  - [4] Disi, L. Gravitation and quantum-mechanical localization of macro-objects. *Phys. Lett. A* 105, 199202. (doi:10.1016/0375-9601(84)90397-9) (1984)
  - [5] G.C. Ghirardi, A. Rimini, and T. Weber. Unified dynamics for microscopic and macroscopic systems. *Phys. Rev. D*, 34:470, (1986).
  - [6] Hawking, S. W. *Nature* **248**, 5443, (1974).



- [7] L. Susskind, L. Thorlacius, J. Uglum *Phys.Rev.D***48**, 3743, (1993).
- [8] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, *JHEP* **02**, 062, (2013).
- [9] O’Connell, A.D., Hofheinz, M., Ansmann, M., Bialczak, R.C., Lenander, M., Lucero, E., Neeley, M., Sank, D., Wang, H., Weides, M., Wenner, J., Martinis, J.M., and Cleland, A.N. *Nature* **464**, 697 (2010).
- [10] Marshall, W., Simon, C., Penrose, R., and Bouwmeester, D. Towards quantum superpositions of a mirror *Phys. Rev. Lett.* **91**, 130401 (2003).
- [11] J. van Wezel, T.H. Oosterkamp, A nanoscale experiment measuring gravity’s role in breaking the unitarity of quantum dynamics. *Proc. R. Soc. A*, 468 (2012) 35-56, <http://dx.doi.org/10.1098/rspa.2011.0201>.
- [12] G. ’t Hooft, Private communications.
- [13] I. Pikovski, M. Zych, F. Costa, C. Brukner, Universal decoherence due to gravitational time dilation, arXiv:1311.1095
- [14] C. Anastopoulos, Quantum theory of nonrelativistic particles interacting with gravity, *Phys. Rev. D* **54**, 1600 (1996).
- [15] M. Zych, F. Costa, I. Pikovski, T.C. Ralph and C. Brukner, General relativistic effects in quantum interference of photons, *Class. Quantum Grav.* **29** 224010 (2012).
- [16] C. Lammerzahl, A Hamilton Operator for quantum optics in gravitational-fields, *Physics letters. A*, **203**(1), 1995, pp. 12-17
- [17] Dimopoulos et al., Testing General Relativity with Atom Interferometry, *Phys. Rev. Letters*, **98**, 111102, (2007).
- [18] Handbook on the Physics and Chemistry of Rare Earths, edited by K. A. Gschneider, Jr. and L. R. Eyring (North-Holland, Amsterdam, 1978); in particular, D. G. Koskenmaki and K. A. Gschneider, Jr., *ibid.*, p. 337.
- [19] Held, K., McMahan, A. K., and Scalettar, R. T. Cerium Volume Collapse: Results from the Merger of Dynamical Mean-Field Theory and Local Density Approximation *Phys. Rev. Lett.* **87**, 276404 (2001).
- [20] Irwin I. Shapiro (1964). "Fourth Test of General Relativity". *Physical Review Letters* **13** (26): 789-791. doi:10.1103/PhysRevLett.13.789